SISMEC Società Italiana di Statistica Medica ed Epidemiologia Clinica

vi congresso nazionale

MISURARE PER MIGLIORARE

ATTI
a cura di
Rosaria Gesuita, Edlira Skrami, Francesca Rosati, Luigi Ferrante
Flavia Carle

Ancona
28 settembre – 1 ottobre 2011
Facoltà di Medicina e Chirurgia

15.20 **USE OF LEXIS DIAGRAMS AND MULTIPLE TIMESCALES IN STEM CELLS TRANSPLANTATION STUDIES** *122*

Iacobelli Simona, Carstensen Bendix

15.40 **Pausa caffè**

Area Metodi Biostatistici

Sessione 2

Introduce e modera: ***Luigi Ferrante,*** *Università Politecnica delle Marche*

16.00 **BAYESIAN MODELING OF CARIES ONSET AND PROGRESSION: THE BELO HORIZONTE CARIES PREVENTION STUDY** *124*

Matranga Domenica, Vullo Angela

16.20 **UN NUOVO MODELLO PARAMETRICO PER IL TEMPO AL PRIMO EVENTO NEL CASO DI EVENTI OVER-DISPERSI: APPLICAZIONE AL TEMPO ALLA PRIMA RICADUTA NELLA SCLEROSI MULTIPLA** *126*

Siri Paola, **Sormani Maria Pia**

16.40 **A STATISTICAL APPROACH TO MULTIPLE RANKING PRIORITIES IN ENVIRONMENTAL EPIDEMIOLOGY** *128*

Catelan Dolores, Biggeri Annibale

17.00 **REAL-TIME PCR QUANTIFICATIONS IN PRESENCE OF INHIBITORS USING SOD AND CY0 METHODS** *130*

Sisti Davide, Guescini Michele, Rocchi Marco BL, Panebianco Renato, Stocchi Vilberto

Auditorium Maria Montessori
sessione plenaria

Ore 17.30 Introduce e modera: ***Stefania Fortuna***, Università Politecnica delle Marche

L'evoluzione della Sanità nei 150 anni dell'Italia Unita ´21

Cosmacini Giorgio, Storico della medicina. Milano

Ore 20.00 **Cena sociale (partenza 19.30)**

BAYESIAN MODELING OF CARIES ONSET AND PROGRESSION: THE BELO HORIZONTE CARIES PREVENTION STUDY

Matranga Domenica[1], Vullo Angela[2]

[1]*Dipartimento di Biopatologia e Biotecnologie Mediche e Forensi*, [2]*Centro Interdipartimentale per il Monitoraggio dell'Economia e del Territorio, Università degli Studi di Palermo*

Introduction. The dmft-index, defined as the number of deciduous teeth (t) that are decayed (d), missing (m) or filled (f) because of caries, is an important indicator of oral health status. In the literature, the relation between the dmft-index and a properly chosen covariate set, is modeled through overdispersed and zero-inflated (ZI) Poisson and Binomial models. In addition, generic mixture models have been proposed as an alternative to zero-inflated models to capture different underlying mechanism of data generation (Giltorpe, 2009). In the literature, Bayesian ZI models exist only for the Poisson distribution (Gosh, 2006; Karlis and Ntzoufras, 2006).

Objectives. The aim of this paper is to propose Bayesian ZI modeling of the dmft-index. In particular, we consider binomial models because the dmft count ranges between 0 and the number of teeth (20,24,28 for deciduous, mixed and permanent respectively). In ZI models, we can think caries onset and progression as estimated through the inflation and the count part, respectively. According to the Bayesian paradigm, the posterior distribution of the parameters represents our uncertainty, updated by data, about the parameters values. It is proportional to the product of the prior distribution and the conditional distribution of the data and all the inference is made on the basis of the posterior distribution on the parameters given the data. If the prior uncertainty is large, the prior distribution is vague. At the opposite, if prior hypotheses about disease progression and onset are available, the prior distribution can be constructed accordingly. The proposed method is applied to the Belo Horizonte CAries Prevention (BELCAP) study concerning 797 brazilian schoolchildren aged 7-years old at the beginning of the study (Böhning, 1999), where only the eight deciduous molars were considered.

Methods. The joint posterior distribution of the parameters of the proposed models turn out to be analytically intractable, hence we performed Markov Chain Monte Carlo (MCMC) simulations to obtain the point and interval estimates of the parameters. We wrote a WinBUGS code to model Binomial, ZI-Binomial and ZI-Beta-binomial models with and without covariates. Models without covariates were implemented through the Gibbs sampling-type algorithms, models with covariates were implemented through the Metropolis-Hasting algorithm. The *logit* link was used to model binomial distribution parameter and the inflation mixture proportion p. In a first attempt we used vague priors to express our lack of knowledge about the regression parameters. Then, we used Normal priors for location parameters of the count and the inflation parts of the ZI-model, with hyperparameters chosen on the base of information available at the WHO web site (http://www.whocollab.od.mah.se/amro/brazil/data/brazilcar.html) for 2002-2003 years. To check model fit, we used posterior mean of the deviance, defined as the negative of twice the log-likelihood function. Parameter estimates were obtained after a burn-in of 5,000 samples with three chains and then every 5th sample was kept, updates followed by further 10,000 updates.

Results. From WHO statistics, the percentage of caries free for 5-years-old children is p=40.7% and the average dmft=2.8. Therefore, the prior was chosen as Normal (logit (0.407), 0.01) for disease onset and as Normal (logit (2.8/8), 0.01) for disease progression. As the considered data-set is large sized, results (see Table) overlap with those obtained with vague priors (data not shown).

Table - Bayesian binomial-type regression models for BELCAP study: coefficients (standard deviations). Priors based on WHO statistics

		ZIB			ZI-Betabinomial	
	Binomial	Constant inflation	Variable inflation	Betabinomial	Constant inflation	Variable inflation
Location	-1.00 (0.08)	-0.70 (0.10)	-0.72 (0.10)	-1.00 (0.12)	-0.88 (0.14)	-1.02 (0.1)
Gender						
Male	0.17 (0.06)	0.13 (0.07)	0.12 (0.07)	0.17 (0.09)	0.17 (0.09)	0.17 (0.08)
Ethnicity						
White	0.13 (0.07)	0.12 (0.08)	0.12 (0.08)	0.12 (0.09)	0.13 (0.10)	0.12 (0.09)
Black	-0.18 (0.10)	-0.14 (0.12)	-0.08 (0.12)	-0.20 (0.14)	-0.21 (0.15)	-0.19 (0.14)
Ethnicity						
Education	-0.32 (0.10)	-0.31 (0.11)	-0.35 (0.12)	-0.28 (0.14)	-0.30 (0.16)	-0.26 (0.13)
All	-0.76 (0.11)	-0.58 (0.13)	-0.47 (0.13)	-0.78 (0.16)	-0.78 (0.16)	-0.76 (0.14)
Enrichment	-0.13 (0.10)	-0.08 (0.11)	-0.08 (0.11)	-0.13 (0.14)	-0.12 (0.14)	-0.11 (0.13)
Rinsing	-0.47 (0.10)	-0.25 (0.11)	-0.18 (0.11)	-0.50 (0.14)	-0.47 (0.15)	-0.49 (0.13)
Hygiene	-0.40 (0.10)	-0.29 (0.12)	-0.26 (0.12)	-0.43 (0.14)	-0.43 (0.15)	-0.43 (0.14)
Over-dispersion				0.17 (0.02)	0.13 (0.03)	0.17 (0.02)
Inflate Location		-1.17 (0.10)	-1.63 (0.32)		-3.44 (3.6)§	-27.9 (20.4)§
Deviance	3103	2209	2198	2151	2095	2151
Estimated % of zeros	14.0%	28.6%	28.7%	27.6%	28.5%	27.6%

MC Error<0.001 except when § used

Six models give very similar parameters estimates. Looking at the deviance criterion, ZI-Betabinomial model with constant inflation should be preferred. However, MC error for the inflate location parameter is higher than 0.001, so this model must be rejected. Betabinomial model is that one with the minimum deviance but its posterior mean of Pr(dmft=0) is the farthest away from the empirical percentage of zero-dmft, which is 29.0%. We conclude that the best model is the ZIB model with variable inflation.

Conclusions. Our analysis is comparable with previous work of Giltorpe et al. (2009) because models setting is the same. We concur with them that ZIB model with variable inflation is the best according to the absolute number of zeros, but we strengthen this result through the deviance criterion. While they suggest latent class modelling for the onset and progression of dmft, we propose Bayesian modelling as a means to take over prior information concerning the mechanism of data generation.

References

Giltorpe MS, Frydenberg M, Cheng Y, Baelum V. *Stat Med* 2009; **28**: 3539-3553
Gosh SK, Mukhopadhyay P, Lu JC. *J Stat Plan Infer* 2006; **136**: 1360-1375
Karlis D, Ntzoufras I. *Stat Med* 2006; 25: 1885-1905
Böhning D, Dietz E, Schlattmann P, Mendoca L, Kirchner U. JRSS Serie A 1999; 162(2): 195-209.